FIG. 1

```
agcttgcggc cgctatttag ggcctttta gatagatgat gcgtttatct acaattagta      60

taaaattagc tttggtggtg aaattaacca ttatagcgtg aaacaaaaag aaaagtaaat     120

atcactggag gtaaacaaaa agaaaggtaa agaaaggctc ctaatcagat gaaaatttct     180

ctcaaacttt tagcactctt ttggaacgta aacagtgaag tgaactagag atgttaaaag     240

aggaagataa atggttaaat atacaaaatg tgtttaacat aaatccattt atcccccatt     300

aattcaatcc tttgaattgt aaatatata atatatgttt aatgtcttcc tgttttaacc     360

atggtaaagg agtcttcttg atttgttaat taagtgaaag gcaccaaaac caaatttata     420

gcgtttaaat tgagtacatt ttgtatatac ggtcgaaagc cgcctcaggg ttcaatggct     480

gctaatattt gcccaaaata aaatctcctt taggacttag cttcgtgtac gatagccaat     540

tcgatatctt atgtaggtaa ctatgtttcg taatacatag aattgttcaa ctttttttat     600

gttcgaactt ttttttaatt tgcaatgtaa ttatttgtct tatgttacat tttagtcatt     660

tatgtttaaa atgttatatt ttagttattt acgttatcgt gttgtaacat tttaggtact     720

gagccattaa ttgtcgttaa cagtgtaaca gaaagctgac gtggcacgtt aaatcatcgc     780

ttcaaattaa aattttggat taaattatac aattggtccc tatatttttt tcattttttt     840

ctttattttt cattctcttc ggcttctccc tatgttttcc tctcttcttc atttctttta     900

acatagtttt tttatatttt ctaaaattaa attttcaaa aaaaataaaa tataggagtt     960

agttttaaaa aatacgttaa aagaaatgga taagggagga aaacagaggg aaaagcagaa    1020

gaggatgaaa gaaaagaaa gttaaaagaa cataaaagaa aaaattaaat tgctctaaaa    1080

gaaaaaatat ggggacagat tgtataaatt agccaaaaat ttttgtttaa aatgatgatt    1140

taagtgctat gtcagtttac cattaaaggt aacaattaac gttttaataa ttaaaatatt    1200

ataacccgat aaaataagta actaaaattt taaattttaa atataaataa ctaaaatata    1260

acttgagata aataaaagtt gctattaaaa tttaggtata tcaacattaa taatgttggc    1320

cccatataag aagaattcat gtgcaagaag ttaaaatggg tcaacagccg cctacatgcc    1380

aacccttaat ttccaacttt tagttggtga ggataagatg agactaaaat ccgacattgc    1440

atattgaaag gtcaagcata atcagataaa aaatcttggt tcattcattg agtaccaagt    1500

aagtcaccct tttttaatca aaactacaac agtgaaggaa actacgaata atctatcatt    1560

caccatgact ttctcccacc atatatttat gtacaccctt ccaatcttcc tacactacca    1620
```

FIG. 2A

```
catcgataaa taaaaactgc agcccggggg atccatagtg taaaaaattc ata atg          1676
                                                              Met
                                                              1


gaa gtc tgc aat tgt att gaa ccg caa tgg cca gcg gat gaa ttg tta          1724
Glu Val Cys Asn Cys Ile Glu Pro Gln Trp Pro Ala Asp Glu Leu Leu
          5                   10                  15


atg aaa tac caa tac atc tcc gat ttc ttc att gcg att gcg tat ttt          1772
Met Lys Tyr Gln Tyr Ile Ser Asp Phe Phe Ile Ala Ile Ala Tyr Phe
          20                  25                  30


tcg att cct ctt gag ttg att tac ttt gtg aag aaa tca gcc gtg ttt          1820
Ser Ile Pro Leu Glu Leu Ile Tyr Phe Val Lys Lys Ser Ala Val Phe
          35                  40                  45


ccg tat aga tgg gta ctt gtt cag ttt ggt gct ttt atc gtt ctt tat          1868
Pro Tyr Arg Trp Val Leu Val Gln Phe Gly Ala Phe Ile Val Leu Tyr
50                   55                  60                   65


gga gca act cat ctt att aac tta tgg act ttc act acg cat tcg aga          1916
Gly Ala Thr His Leu Ile Asn Leu Trp Thr Phe Thr Thr His Ser Arg
                  70                  75                  80


acc gtg gcg ctt gtg atg act acc gcg aag gtg tta acc gct gtt gtc          1964
Thr Val Ala Leu Val Met Thr Thr Ala Lys Val Leu Thr Ala Val Val
              85                  90                  95


tcg tgt gct act gcg ttg atg ctt gtt cat att att cct gat ctt ttg          2012
Ser Cys Ala Thr Ala Leu Met Leu Val His Ile Ile Pro Asp Leu Leu
              100                 105                 110


agt gtt aag act cgg gag ctt ttc ttg aaa aat aaa gct gct gag ctc          2060
Ser Val Lys Thr Arg Glu Leu Phe Leu Lys Asn Lys Ala Ala Glu Leu
          115                 120                 125


gat aga gaa atg gga ttg att cga act cag gaa gaa acc gga agg cat          2108
Asp Arg Glu Met Gly Leu Ile Arg Thr Gln Glu Glu Thr Gly Arg His
130                 135                 140                 145


gtg aga atg ttg act cat gag att aga agc act tta gat aga cat act          2156
Val Arg Met Leu Thr His Glu Ile Arg Ser Thr Leu Asp Arg His Thr
                  150                 155                 160


att tta aag act aca ctt gtt gag ctt ggt agg aca tta gct ttg gag          2204
Ile Leu Lys Thr Thr Leu Val Glu Leu Gly Arg Thr Leu Ala Leu Glu
              165                 170                 175


gag tgt gca ttg tgg atg cct act aga act ggg tta gag cta cag ctt          2252
Glu Cys Ala Leu Trp Met Pro Thr Arg Thr Gly Leu Glu Leu Gln Leu
          180                 185                 190
```

FG. 2B

```
tct tat aca ctt cgt cat caa cat ccc gtg gag tat acg gtt cct att          2300
Ser Tyr Thr Leu Arg His Gln His Pro Val Glu Tyr Thr Val Pro Ile
    195                 200                 205


caa tta ccg gtg att aac caa gtg ttt ggt act agt agg gct gta aaa          2348
Gln Leu Pro Val Ile Asn Gln Val Phe Gly Thr Ser Arg Ala Val Lys
210                 215                 220                 225


ata tct cct aat tct cct gtg gct agg ttg aga cct gtt tct ggg aaa          2396
Ile Ser Pro Asn Ser Pro Val Ala Arg Leu Arg Pro Val Ser Gly Lys
            230                 235                 240


tat atg cta ggg gag gtg gtc gct gtg agg gtt ccg ctt ctc cac ctt         2444
Tyr Met Leu Gly Glu Val Val Ala Val Arg Val Pro Leu Leu His Leu
            245                 250                 255


tct aat ttt cag att aat gac tgg cct gag ctt tca aca aag aga tat         2492
Ser Asn Phe Gln Ile Asn Asp Trp Pro Glu Leu Ser Thr Lys Arg Tyr
            260                 265                 270


gct ttg atg gtt ttg atg ctt cct tca gat agt gca agg caa tgg cat        2540
Ala Leu Met Val Leu Met Leu Pro Ser Asp Ser Ala Arg Gln Trp His
            275                 280                 285


gtc cat gag ttg gaa ctc gtt gaa gtc gtc gct gat cag gtg gct gta        2588
Val His Glu Leu Glu Leu Val Glu Val Val Ala Asp Gln Val Ala Val
290                 295                 300                 305


gct ctc tca cat gct gcg atc cta gaa gag tcg atg cga gct agg gac        2636
Ala Leu Ser His Ala Ala Ile Leu Glu Glu Ser Met Arg Ala Arg Asp
            310                 315                 320


ctt ctc atg gag cag aat gtt gct ctt gat cta gct aga cga gaa gca        2684
Leu Leu Met Glu Gln Asn Val Ala Leu Asp Leu Ala Arg Arg Glu Ala
            325                 330                 335


gaa aca gca atc cgt gcc cgc aat gat ttc cta gcg gtt atg aac cat        2732
Glu Thr Ala Ile Arg Ala Arg Asn Asp Phe Leu Ala Val Met Asn His
            340                 345                 350


gaa atg cga aca ccg atg cat gcg att att gca ctc tct tcc tta ctc        2780
Glu Met Arg Thr Pro Met His Ala Ile Ile Ala Leu Ser Ser Leu Leu
            355                 360                 365


caa gaa acg gaa cta acc cct gaa caa aga ctg atg gtg gaa aca ata        2828
Gln Glu Thr Glu Leu Thr Pro Glu Gln Arg Leu Met Val Glu Thr Ile
370                 375                 380                 385


ctt aaa agt agt aac ctt ttg gca act ttg atg aat gat gtc tta gat        2876
Leu Lys Ser Ser Asn Leu Leu Ala Thr Leu Met Asn Asp Val Leu Asp
            390                 395                 400
```

FIG. 2C

```
ctt tca agg tta gaa gat gga agt ctt caa ctt gaa ctt ggg aca ttc      2924
Leu Ser Arg Leu Glu Asp Gly Ser Leu Gln Leu Glu Leu Gly Thr Phe
            405                 410                 415

aat ctt cat aca tta ttt aga gag gtc ctc aat ctg ata aag cct ata      2972
Asn Leu His Thr Leu Phe Arg Glu Val Leu Asn Leu Ile Lys Pro Ile
            420                 425                 430

gcg gtt gtt aag aaa tta ccc atc aca cta aat ctt gca cca gat ttg      3020
Ala Val Val Lys Lys Leu Pro Ile Thr Leu Asn Leu Ala Pro Asp Leu
            435                 440                 445

cca gaa ttt gtt gtt ggg gat gag aaa cgg cta atg cag ata ata tta      3068
Pro Glu Phe Val Val Gly Asp Glu Lys Arg Leu Met Gln Ile Ile Leu
450                 455                 460                 465

aat ata gtt ggt aat gct gtg aaa ttc tcc aaa caa ggt agt atc tcc      3116
Asn Ile Val Gly Asn Ala Val Lys Phe Ser Lys Gln Gly Ser Ile Ser
                470                 475                 480

gta acc gct ctt gtc acc aag tca gac aca cga gct gct gac ttt ttt      3164
Val Thr Ala Leu Val Thr Lys Ser Asp Thr Arg Ala Ala Asp Phe Phe
                485                 490                 495

gtc gtg cca act ggg agt cat ttc tac ttg aga gtg aag gta aaa gac      3212
Val Val Pro Thr Gly Ser His Phe Tyr Leu Arg Val Lys Val Lys Asp
            500                 505                 510

tct gga gca gga ata aat cct caa gac att cca aag att ttc act aaa      3260
Ser Gly Ala Gly Ile Asn Pro Gln Asp Ile Pro Lys Ile Phe Thr Lys
            515                 520                 525

ttt gct caa aca caa tct tta gcg acg aga agc tcg ggt ggt agt ggg      3308
Phe Ala Gln Thr Gln Ser Leu Ala Thr Arg Ser Ser Gly Gly Ser Gly
530                 535                 540                 545

ctt ggc ctc gcc atc tcc aag agg ttt gtg aat ctg atg gag ggt aac      3356
Leu Gly Leu Ala Ile Ser Lys Arg Phe Val Asn Leu Met Glu Gly Asn
                550                 555                 560

att tgg att gag agc gat ggt ctt gga aaa gga tgc acg gct atc ttt      3404
Ile Trp Ile Glu Ser Asp Gly Leu Gly Lys Gly Cys Thr Ala Ile Phe
            565                 570                 575

gat gtt aaa ctt ggg atc tca gaa cgt tca aac gaa tct aaa cag tcg      3452
Asp Val Lys Leu Gly Ile Ser Glu Arg Ser Asn Glu Ser Lys Gln Ser
            580                 585                 590
```

FIG. 2D

```
ggc ata ccg aaa gtt cca gcc att ccc cga cat tca aat ttc act gga        3500
Gly Ile Pro Lys Val Pro Ala Ile Pro Arg His Ser Asn Phe Thr Gly
    595             600             605

ctt aag gtt ctt gtc atg gat gag aac ggg gta agt aga atg gtg acg        3548
Leu Lys Val Leu Val Met Asp Glu Asn Gly Val Ser Arg Met Val Thr
610             615             620             625

aag gga ctt ctt gta cac ctt ggg tgc gaa gtg acc acg gtg agt tca        3596
Lys Gly Leu Leu Val His Leu Gly Cys Glu Val Thr Thr Val Ser Ser
                630             635             640

aac gag gag tgt ctc cga gtt gtg tcc cat gag cac aaa gtg gtc ttc        3644
Asn Glu Glu Cys Leu Arg Val Val Ser His Glu His Lys Val Val Phe
        645             650             655

atg gac gtg tgc atg ccc ggg gtc gaa aac tac caa atc gct ctc cgt        3692
Met Asp Val Cys Met Pro Gly Val Glu Asn Tyr Gln Ile Ala Leu Arg
        660             665             670

att cac gag aaa ttc aca aaa caa cgc cac caa cgg cca cta ctt gtg        3740
Ile His Glu Lys Phe Thr Lys Gln Arg His Gln Arg Pro Leu Leu Val
        675             680             685

gca ctc agt ggt aac act gac aaa tcc aca aaa gag aaa tgc atg agc        3788
Ala Leu Ser Gly Asn Thr Asp Lys Ser Thr Lys Glu Lys Cys Met Ser
690             695             700             705

ttt ggt cta gac ggt gtg ttg ctc aaa ccc gta tca cta gac aac ata        3836
Phe Gly Leu Asp Gly Val Leu Leu Lys Pro Val Ser Leu Asp Asn Ile
                710             715             720

aga gat gtt ctg tct gat ctt ctc gag ccc cgg gta ctg tac gag ggc        3884
Arg Asp Val Leu Ser Asp Leu Leu Glu Pro Arg Val Leu Tyr Glu Gly
        725             730             735

atg taa ggatccagct ttcgttcgta tcatcggttt cgacaacgtt cgtcaagttc        3940
Met

aatgcatcag tttcattgcg cacacaccag aatcctactg agttcgagta ttatggcatt        4000

gggaaaactg tttttcttgt accatttgtt gtgcttgtaa tttactgtgt tttttattcg        4060

gttttcgcta tcgaactgtg aaatggaaat ggatggagaa gagttaatga atgatatggt        4120

ccttttgttc attctcaaat taatattatt tgttttttct cttatttgtt gtgtgttgaa        4180

tttgaaatta taagagatat gcaaacattt tgttttgagt aaaaatgtgt caaatcgtgg        4240
```

FIG. 2E

```
cctctaatga ccgaagttaa tatgaggagt aaaacacttg tagttgtacc attatgctta    4300

ttcactaggc aacaaatata ttttcagacc tagaaaagct gcaaatgtta ctgaatacaa    4360

gtatgtcctc ttgtgtttta gacatttatg aactttcctt tatgtaattt tccagaatcc    4420

ttgtcagatt ctaatcattg ctttataatt atagttatac tcatggattt gtagttgagt    4480

atgaaaatat tttttaatgc attttatgac ttgccaattg attgacaaca tgcatcaatc    4540

gacctgcagc cactcgaagc ggccgccact cgagtggaag ctagcttccc gatcctatct    4600

gtcacttcat caaaaggaca gtagaaaagg aaggtggcac tacaaatgcc atcattgcga    4660

taaaggaaag gctatcgttc aagatgcctc tgccgacagt ggtcccaaag atggaccccc    4720

acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa agcaagtgga    4780

ttgatgtgat acttccactg acgtaaggga tgacgcacaa tcccactatc cttcgcaaga    4840

cccttcctct atataaggaa gttcatttca tttggagagg acacgctgaa atcaccagtc    4900

tctctctaca agatcgggga tctctagcta gacgatcgtt tcgcatgatt gaacaagatg    4960

gattgcacgc aggttctccg gccgcttggg tggagaggct attcggctat gactgggcac    5020

aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag gggcgcccgg    5080

ttcttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac gaggcagcgc    5140

ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg    5200

aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc ctgtcatctc    5260

accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg ctgcatacgc    5320

ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta    5380

ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat caggggctcg    5440

cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag gatctcgtcg    5500

tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc ttttctggat    5560

tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg ttggctaccc    5620

gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg ctttacggta    5680

tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct tcttgacgag ttcttctgag    5740

cgggactctg gggttcgatc cccaattccc gatcgttcaa acatttggca ataaagtttc    5800
```

FIG. 2F

```
ttaagattga atcctgttgc cggtcttgcg atgattatca tataatttct gttgaattac    5860

gttaagcatg taataattaa catgtaatgc atgacgttat ttatgagatg ggttttatg     5920

attagagtcc cgcaattata catttaatac gcgatagaaa acaaaatata gcgcgcaaac    5980

taggataaat tatcgcgcgc ggtgtcatct atgttactag atcggggatc gggccactcg    6040

agtggtggcc gcatcgatcg tgaagtttct catctaagcc cccatttgga cgtgaatgta    6100

gacacgtcga aataaagatt tccgaattag aataatttgt ttattgcttt cgcctataaa    6160

tacgacggat cgtaatttgt cgttttatca aaatgtactt tcattttata ataacgctgc    6220

ggacatctac attttttgaat tgaaaaaaaa ttggtaatta ctctttcttt ttctccatat   6280

tgaccatcat actcattgct gatccatgta gatttcccgg acatgaagcc atttacaatt    6340

gaatatatcc tgccgccgct gccgctttgc acccggtgga gcttgcatgt tggtttctac    6400

gcagaactga gccggttagg cagataattt ccattgagaa ctgagccatg tgcaccttcc    6460

ccccaacacg gtgagcgacg gggcaacgga gtgatccaca tgggactttt cctagcttgg    6520

ctgccatttt tggggtgagg ccgttcgcgc ggggcgccag ctggggggat gggaggcccg     6580

cgttaccggg agggttcgag aagggggggc acccccttc ggcgtgcgcg gtcacgcgcc     6640

agggcgcagc cctggttaaa aacaaggttt ataaatattg gtttaaaagc aggttaaaag    6700

acaggttagc ggtggccgaa aaacgggcgg aaacccttgc aaatgctgga ttttctgcct     6760

gtggacagcc cctcaaatgt caataggtgc gcccctcatc tgtcatcact ctgcccctca     6820

agtgtcaagg atcgcgcccc tcatctgtca gtagtcgcgc ccctcaagtg tcaataccgc     6880

agggcactta tccccaggct tgtccacatc atctgtggga aactcgcgta aaatcaggcg     6940

ttttcgccga tttgcgaggc tggccagctc cacgtcgccg gccgaaatcg agcctgcccc    7000

tcatctgtca acgccgcgcc gggtgagtcg gccccctcaag tgtcaacgtc cgcccctcat    7060

ctgtcagtga gggccaagtt ttccgcgtgg tatccacaac gccggcggcc ggccgcggtg     7120

tctcgcacac ggcttcgacg gcgtttctgg cgcgtttgca gggccataga cggccgccag     7180

cccagcggcg agggcaacca gcccggtgag cgtcggaaag ggtcgatcga ccgatgccct     7240

tgagagcctt caacccagtc agctccttcc ggtgggcgcg gggcatgact atcgtcgccg      7300

cacttatgac tgtcttcttt atcatgcaac tcgtaggaca ggtgccggca gcgctctggg     7360
```

FIG. 2G

```
tcattttcgg cgaggaccgc tttcgctgga gcgcgacgat gatcggcctg tcgcttgcgg    7420

tacgccctcg ctcaagcctt cgtcactggt cccgccacca aacgtttcgg cgagaagcag    7480

gccattatcg ccggcatggc ggccgacgcg ctgggctacg tcttgctggc gttcgcgacg    7540

cgaggctgga tggccttccc cattatgatt cttctcgctt ccggcggcat cgggatgccc    7600

gcgttgcagg ccatgctgtc caggcaggta gatgacgacc atcagggaca gcttcaagga    7660

tcgctcgcgg ctcttaccag cctaacttcg atcactggac cgctgatcgt cacggcgatt    7720

tatgccgcct cggcgagcac atggaacggg ttggcatgga ttgtaggcgc cgccctatac    7780

cttgtctgcc tccccgcgtt gcgtcgcggt gcatggagcc gggccacctc gacctgaatg    7840

gaagccggcg gcacctcgct aacggattca ccactccaag aattggagcc aatcaattct    7900

tgcggagaac tgtgaatgcg caaaccaacc cttggcagaa catatccatc gcgtccgcca    7960

tctccagcag ccgcacgcgg cgcatctcgg gcagcgttgg gtcctggcca cgggtgcgca    8020

tgatcgtgct cctgtcgttg aggacccggc taggctggcg gggttgcctt actggttagc    8080

agaatgaatc accgatacgc gagcgaacgt gaagcgactg ctgctgcaaa cgtctgcga    8140

cctgagcaac aacatgaatg gtcttcggtt ccgtgtttc gtaaagtctg gaaacgcgga    8200

agtcagcgcc ctgcaccatt atgttccgga tctgcatcgc aggatgctgc tggctaccct    8260

gtggaacacc tacatctgta ttaacgaagc gctggcattg accctgagtg atttttctct    8320

ggtcccgccg catccatacc gccagttgtt taccctcaca acgttccagt aaccgggcat    8380

gttcatcatc agtaacccgt atcgtgagca tcctctctcg tttcatcggt atcattaccc    8440

ccatgaacag aaattccccc ttacacggag gcatcaagtg accaaacagg aaaaaaccgc    8500

ccttaacatg gcccgcttta tcagaagcca gacattaacg cttctggaga aactcaacga    8560

gctggacgcg gatgaacagg cagacatctg tgaatcgctt cacgaccacg ctgatgagct    8620

ttaccgcagc tgcctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct    8680

cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg    8740

cgcgtcagcg ggtgttggcg ggtgtcgggg cgcagccatg acccagtcac gtagcgatag    8800

cggagtgtat actggcttaa ctatgcggca tcagagcaga ttgtactgag agtgcaccat    8860

atgcggtgtg aaataccgca cagatgcgta aggagaaaat accgcatcag gcgctcttcc    8920

gcttcctcgc tcactgactc gctgcgctcg tcgttcggc tgcggcgagc ggtatcagct    8980
```

FIG. 2H

```
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg   9040

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc   9100

cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga   9160

aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct   9220

cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg   9280

gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag   9340

ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat   9400

cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac   9460

aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac   9520

tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc   9580

ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt   9640

tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc   9700

ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg   9760

agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca   9820

atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca   9880

cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag   9940

ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac   10000

ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc   10060

agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct   10120

agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tgcaggtcgg   10180

gagcacagga tgacgcctaa caattcattc aagccgacac cgcttcgcgg cgcggcttaa   10240

ttcaggagtt aaacatcatg agggaagcgg tgatcgccga agtatcgact caactatcag   10300

aggtagttgg cgtcatcgag cgccatctcg aaccgacgtt ctggccgta catttgtacg   10360

gctccgcagt ggatggcggc ctgaagccac acagtgatat tgatttgctg gttacggtga   10420

ccgtaaggct tgatgaaaca acgcggcgag ctttgatcaa cgacctttg gaaacttcgg   10480

cttcccctgg agagagcgag attctccgcg ctgtagaagt caccattgtt gtgcacgacg   10540

acatcattcc gtggcgttat ccagctaagc gcgaactgca atttggagaa tggcagcgca   10600
```

FIG. 2I

```
atgacattct tgcaggtatc ttcgagccag ccacgatcga cattgatctg gctatcttgc    10660

tgacaaaagc aagagaacat agcgttgcct tggtaggtcc agcggcggag gaactctttg    10720

atccggttcc tgaacaggat ctatttgagg cgctaaatga aaccttaacg ctatggaact    10780

cgccgcccga ctgggctggc gatgagcgaa atgtagtgct tacgttgtcc cgcatttggt    10840

acagcgcagt aaccggcaaa atcgcgccga aggatgtcgc tgccgactgg gcaatggagc    10900

gcctgccggc ccagtatcag cccgtcatac ttgaagctag gcaggcttat cttggacaag    10960

aagatcgctt ggcctcgcgc gcagatcagt tggaagaatt tgttcactac gtgaaaggcg    11020

agatcaccaa ggtagtcggc aaataatgtc taacaattcg ttcaagccga cgccgcttcg    11080

cggcgcggct taactcaagc gttagatgct gcaggcatcg tggtgtcacg ctcgtcgttt    11140

ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg    11200

ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg aggatttttc ggcgctgcgc    11260

tacgtccgck accgcgttga gggatcaagc cacagcagcc cactcgacct ctagccgacc    11320

cagacgagcc aagggatctt tttggaatgc tgctccgtcg tcaggctttc cgacgtttgg    11380

gtggttgaac agaagtcatt atcgtacgga atgccaagca ctcccgaggg gaaccctgtg    11440

gttggcatgc acatacaaat ggacgaacgg ataaaccttt tcacgccctt ttaaatatcc    11500

gttattctaa taaacgctct tttctcttag gtttacccgc caatatatcc tgtcaaacac    11560

tgatagttta aactgaaggc gggaaacgac aatctgatcc ccatca                   11606
```

FIG. 2J

FIG. 3A

NotI

AGCTTGCGCCGCGCTATTTAGGGCCTTTTTAGATAGATGATGCGTTTATCTACAATTAGTATAAAAATTAGCTTTGGTGTGTGAAATTAACCATTATAGCGTGAAACAAAAGAAAA
GTAAATATCACTGGAGGTAAACAAAAGAAAAGAAGGCTAAGAAAGTAAGAAATTTCTCTCAATCAGATGAAAATTTCTCCAATCTTTTGAACACTCTTTGGAACACAGTGAACTAG
AGATGTTAAAAGAGGAAGAATAAATGTTTAAATATACAAAATGTGTTTAACATAAATGTTTAACATAAATCCATTAATCAAATCCTTGAATTGTTAAATATATAATATATGTTTAA

NcoI

TGTCTTCCTGTTTAACCATGGTAAGGAGTCTTCTTGATTTGTTAATTAAGTGAAAGGCACCAAAACCAAAATTTATAGCGTTTAAATTGAGTACATTTTGTATATACGGTCGA

EcoRV

AAGCCGCCTCAGGGTTCAATGGCTGCTAATATTTGCCCAAAATAAAAATCTCCTTAGGACTTGCGTGTACGATAGCCAATTCGATATCTTATGTAGGTAACTATGTTCG
TAATACATTAGAATTGTTCAACTTTTTTATGTTCGAACTTTTTTGCAATGTAATTATTTGTCTTATTGTTCATTAGTTCATTTAGTGTTTTAAAAATGTTATATTTAG
TTATTACGTTATCGTGTTGTAACATTTAGTAGTACTGAGCCATTAATTGTCGTTAACAGAAAGCTGACGTGGCACGTTAAATCATCGCTTAAATTCAAATTAAAATTTGG
ATTAAAATTATACAATTGGTCCCTATATTTTTCATTTTTTTATTTTCTTATTTTCATTGTTTTCATTTTTTTCTTCATTTTCTTTTCATTTCTTTTAACATAGTTTTT
TATATTTTCTAAAATTAAATTTTCAAAAAATAAAATATAGGAGTTAGTTTAAAAAATACGTTAAAAATAAGGGACAGATTGTATAAAAATATGGGAAAAGCAGAGAGGAT
GAAAGAAAAGAAAGTTAAAAGAACATAAAGAGAAAAATTAAATTGCTCTAAAGAATTAAAATTTATAATTTTGTTTAAAATGATGATT
TAAGTGCTATGTCAGTTTACCATTAAAGTAAGTAACAATTAACGTTTTAATAAAAATATTATAACCCGATAAAAATAAGTAACTAAAATTTTAAATTTTAAAATATAAAACTAA

EcoRI

AATATAACTTGAGATAAATAAAAAGTTGCTATTAAAAATTTAGGTATATCAACATTAATAAATGTTGGCCCATATAAGAAGAATTCATGTGCAAGAAGTTAAAATGGTCAACAGC
CGCCTACACATGCAACCCTTAATTTCCAACTTTTAGTTGGTGAGGATAAGATGAGACTAAAATCCGACATTGCATAATTGCAAGGTCAAGCATAATCAGATAAAAAATCTTGGTTC
ATTCATTGAGTAGTACCAAGTAAGTCACCCTTTTTTTTAATCAAACCTAGGAAAACTACGAATAATCTATCATTCACCATATAATCTATCATTTATGTACAC

SmaI

CCTTCCAATCTTCCTACACTACCACATCGATAAATAAACTGCAGCCCGGGGATCCATAGTGTAAAAAATTCATAATGGAAGTCTGCAATTGTATTGAACCGCAATGGCCAG
                   ClaI   PstI   SmaI BamHI   >MetGluValCysAsnCysIleGluProGlnTrpProA
CGGATGAATGTTAATGAAATACCAATACACTCCGATTCTTCTTCATTGCGATTGCGTATTGCGATTCCTTGAGTGATTACTTTGTGAAGAAATCAGCCGTGTTTCCGT
laAspGluLeuLeuMetLysTyrGlnTyrIleSerAspPhePheIleAlaIleAlaTyrPhePheSerIleProLeuLeuIleLeuTyrPheValLysSerAlaValPheProT
ATAGATGGGTACTTGTTCAGTTTGGTGTCTTTATCGTTCTTTATGGAGCAACTCATCTTATTAACTTATGGAGTAGCGTTCACTACGCATTCGAGAACCGTGGCGCTTGTGATGACTA
yrArgTrpValLeuValGlnPheGlyAlaPheIleValLeuTyrGlyAlaThrHisLeuLeuAsnLeuTrpPhePheThrThrHisSerArgThrValAlaLeuValMetThrT
CCGCGAAGGTGTTAACCGCTGTTGTCTCGTCTACTGCGTTCATGCGTTCATATTATTCCTGATCTTGTGTTACTACTTTGAGTGTTAAGACTCGGGAGCTTTTCTGAAAAATAAAGCTG
hrAlaAlaLysValThrAlaValSerCysAlaValAlaThrAlaLeuMetLeuValHisIleIleProAspLeuLeuSerValLysThrArgGluLeuPheLeuLysAsnLysAlaA
CTGAGCTCGATAGAGAAATGGGATTGATTCGAACTCAGAAGAAACCGGAAGCATGTGAAGTGTTGACTACTTAGATAGACATACTATTTAAAGA
laGluLeuAspArgGluMetGlyLeuIleLeuIleLeuArgThrGlnGluThrGlyValArgHisValArgMetLeuThrHisGlyIleAspArgSerThrLeuAspArgHisThrIleLeuLysT
CTACACTTGTTGAGCTTGAGCTTTGGTGGTGTTGGATGTTGGAGGAGTGCCTACTAGCTTGGTTAGAGCTACAGCTTTCTTATACACTTCGTCATCCAACATCCCG
hrThrLeuValGluLeuGluLeuGlyLeuGluLeuLeuValGluGluCysAlaLeuAlaLeuThrArgTyrThrLeuArgHisGlnHisProV

SpeI

TGGAGTATACGGTTCCTATTCAATTACCGGTGATTAACCAAGTGTTTGGTACTGATTAGGGCTGTAAAAATATCTCCTAATTCTCCCTGTGTGGCTAGGTTGAGACCTGTTTCTGGGA
alGluTyrThrValProLeuGlnLeuProValIleLeuAsnGlnValPheGlyThrSerArgAlaValLysIleSerProAsnSerProAlaArgLeuArgProValSerGlyLeuL
AATATATGCTAGGGGAGGTGGTCGCTGCGTGAGGTTCCGTTCCTCCACCTTTCAATTTTCAGATTAATGACTGGCGTAGCTTTCAACAAGATATGCTTTGATGGTTTGA
ysIleTyrMetLeuGlyGlyArgCysArgAlaValAlaValArgValProLeuLeuHisLeuSerLeuSerAsnPheGlnIleLeuAsnAspTrpProGluLeuSerThrLysArgTyrAlaLeuMetValLeuM
TGCTTCCTTCAGATAGTGCAAGGCAATGCAGTCCATGAGTTGAACTCGTTGCGATGTCCATGACTGGCTTCCTCAGTCGCCTCTCACATGCGATCGCGATCCTAGAGAGTCGA

FIG. 3B

BglII

SmaI

BamHI

XhoI

XbaI

SmaI

XhoI

PstI

NotI

TCGTTCAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCACCCACGAGGAGCATCGTGGAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTG
ATACTTCCACTGACGCGTAAGGGATGACGCGCACCAACTCCCACTATCGTTTCGCATGATGAACAAGATGGATTGAACAGAGAGAAGTTCATTTCATTTGGAGAGGACACGCTGAAATCACCCAGTCTC

TCTCTACAAGATCGGGGGGATCTCTAGCTAGACGATCGTTTCGCATGATGATCCTCGTGTTCGGCTGTCCGCTGTCCGGCTGTCCGGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGAC
PstI
GCACAACAGACAATCGGCTGCTCTGATGCCGTTCCGGCTGTCCGGCTGTCCGGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGAC
PvuII

GAGGCAGCGCGGCTATCGTGGCCACGACGGCGTTCCTGCCGAGTGTGCTCGCAGCTGTGCTCGACGTGTCTGCTGCTCTTCCTGCCGAGGCGGGAAGGGACTGCGGCTGTCAGCTGTCACCTGCCGAAGTGCCGGGGCAG
GATCTCCTGTCGTCATCCTGCTCTGCTCCTGCCGAGAAGTATCCATCATCGAAAGTATCCATCATCGGCGGCTGATCGGGCCGGCTTGATCCGGCGCATACGCTTGATCCGGCGGCTGATGCGGCAAAGTATCCATCATCGGCGGCTGATCGGCCGGCTTGATCCGGCGCATACGCTTGATCCGGCGAAA
CATCGCAGCGAGCGAGCTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGCGATCAGGAGGATGATCTCGGCCCAGCCGAACTGTTCGCCCAGCTCAAGGCG
NcoI
CGCATGCCCGACGGCGAGGATCTCGTCGTGACCATGGCCATGGCCGATGCCGCTGCTTGCCGAATATCATGGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTG
GCGGACCGGCTATCAGGACACATAGCGTTGGCTACCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTATCCCGATTCGCAG
PvuI

CGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGGACTCGGGGGTTcgatcccaAATTccgGATCGTTCAAACATTGGCAATAAAGTTTCTTAAGATTGAATCC
TGTTGCCGGTCTTGCGATGATTATCATATATAATTTCTGTTTGAATTACGTTAAGCATGTTAAGCATGTTATTATGACGTTATTTATGATGGTTTTTATGATTAGAGT
XhoI
CCCGCAATTATACATTTAATACGCGATAGAAAAACAAAATATAGCGCCAAACTAGGATAAAATTATCGCGCGCGGTGTCATCCTATGTTACTAGATCGggatcGGGCCACTCGAg
PvuI
ClaI

tggtggccgcatCGATCGTGAAGTTTCTCATCTAAGCCCCCATTTGGACGTGAATGTAGACACGTCGAAATAAAGATTCGAATTAGAATAATTGTGTTATTGCTTCGCCTA
TAAATACGACGGATCGTAATTTGTCGTTTATCGTTTATCAAAATGTACTTTCATTTTCAAAATGTACTTTCATTTTGAATTGAAAAAAAATTGGTAATTACTCTTTCTTT
TTCTCCCATATTGACCATCATTGCTGATTCCCGGACATGTAGATTTCCCGGACATGTAGCATGTGCAGATAAATTTCCGATTCAGGCGTTTTCGCGATTTTG
GCATGTTGGTTTCTTACGGCAGAACTGAGCCGGTTAGGCAGGCAGATAATTTCATTGAGAACTGAGCCCATGTCAGATAAATTTCATTGAGAACTAGAGTATCCA
PvuII

CATGGGACTTTTcctAGCTTGGCTGCCATTTTTGGGGTGAGGCCGTTCGCGCGGCGATGGGAGGCCCGCGTTACCGGAGGGTTCGAGAAGGGGGGCAC
CCCCCTTCGGCGTGGCGCGGTCACGCGCCAGGGCGCAGCCCTGGTTAAAACAAGGTTTATAAATATTGGTTAAAAGACAGGTTAGCGGTGGCCGAAAAACGG
GCGGAAACCCTTGCAAATGCTGGATTTTCTGCCTGTGGACAGCCCTCAAATGTCGCGCCCTCATCTGTCATCACTCTGCCCTCAAGTGTCAAGGATCGCGCCCC
TCATCTGTCAGTAGTCGCCCCCTCAAGTGTCAATACGCCAGGCACTTATCCCCAGGCACTTGTCTGTGGAAACTCGCGTAAATCAGGCGTTTTCGCGATTTTG
CGAGGCTGGCCAGCGCACGTCGCGGCCGAAATCGAGCCGGCCCCCTCATCGTCAAGTGTCGGCCCCTCAAGTGTCAAGCGTCTCCGCCCTCATCGTCAGT
GAGGGCCAAGTTTCCGCGGGTATCCACAACGCCCGGTGTCTCGCCACACGGCCTTCGGGCGCGTTCTGCGGCGCGGCCCATAGACGGCCGGCCAGCC
PvuI

CAGCGGCGAGGGCAACCAGCCCGGTCAGCCGCATCGACCGGTCGAAAGGGTCGATCGACAGCCTTCAACCCAGTCAGCTCCTTCCTTCGCGGTGGCGGGGCATGACTATCGTC
GCCGCACTTGACTGCTCTCTTTATCATGCAAGTCGTAGGACAGAGTGCCGGCCACCTCGGGGTCATTTCGGCGGAGGACCGGCCTTCGGCGGCGGCGACGATGATCGGCCTG
TCGCTTGGCGTACGCCCCTCGTCAGCCTTGCGGGCGCCAAACGTTTCGGGCGGAGAAGCAGGCCATTATCGCCGGCGTTCGGGCGGCTTCGCGGCTACGTCTT
GCTGGCGTTCGGCGAGGCTGGATGGCCTTCCCGGCCCGGCCATGCGGGCCATGCGGGCATCGGCGTTGCAGGCGTTGCAGGTGTCAGGCAGTAGÄTGACGA
CCATCAGGGACAGCTTCAAGGATCGCTCGAAGGATCGCTCTTACCAGCTCACTGAACTTGTCGACCGCTGATTTATGCCGGCGGCGATTTATGCCGCTCGGCGAGCCACATGGAACGGGTT

GGCATGGATTGTAGGGCGGCGCCCTATACCTTGTCTGCCTCCCCGCGTGTTGCGTCGGCGTGCATGGAGGAGCCGGGCGGCCACCTCGACCTCGAATGGAAGCCGGCGGCGGCCACCTCGCTAACGGA
TTCACCACTCCCAAGAATTGGAGCCAATCAATTCTTGCGCAGGAACTGTGAATGCCAAACCACCTTGGCAGAACATATCCGGTCCGCCATCTCCAGCGCCGCACGCGG
CGCATCTCGGGCACGCGTTGGGTCCTGGCCACGGGTGCGCATGAGTCGCCCTGTCGTTGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCG
ATACGGCAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACACCTGAGCAACAACATGAATGGTCTTCGTTTCGTTGGAAACTCTGGAAACGCGGAGTCAGCG
CCCTGCACCATTATGTTCCGGATCTGCATCGCAGGATGCTGCTGCACCCTGTGGAACACCTACCCTGTGAACATCTGTATTAACGGCATGGCTGGAGTGACCCTGAGTGATTTTCTCTGG
TCCCGCCGCATCCATACCGCCAGTTGTTCTTTGTTTACCCTGCACAACGTTCATCATCAGTGTTCATCGCGGTATCGGTATCAACG
ACCCCCATGAACAGAGAAATTCCCCGTTACACGGAGGCATCAAGTGACCAAACAGGAAAAAACCGCCCTTAACATGGCCCGGTTTATCAGAAGCCAGACATTAACGCTTCTGGAGA
PvuII
AACTCAACGAGCTGGACGCGGATGAACAGGCAGCATCTGTGAATCGCTTCACGACCACGCTGCTTCGCGCGTTCGGTGATGACGGTGAAAACC
TCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCGGCAGCAAGCCCGTCAGCAGAGTCCCGGGCGCAGCCA
TGACCCAGTCAGTAGCGGATAGCGGAGTGTATACTGGCCTTAACTATGCCGCTCCTCCGCCTGACTGCGTCGGGTATCAGCGTATACGGTTATC
GAGAAATACCGCCGATCCATCAGGCCGCTCTTCCCGCTCGCGCCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGCGTTATACGGTTATC
CACAGAATCAGCGGGATAACGCCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTTTTCCCATAGGCTCCGCCCCCTGA
CGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCT
GCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTG
TGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
GATTAGCAGAGCGAGGTATGTGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC
CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA
ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGT
TGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAAT
AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAG
PstI
TTTGCGCAACGTTGTTGCCATTGCTGCAGGTCggagcacaggatgacgctaacaattcattcaagccgacacccgcttcgcggcgcggcttaattcaggagttaaacatcatg
aggaagcggtggtgatcgcgccgacgtatatcagaggtagtgttggcgtcatcgactcattggtagcgtcatgcccgacgttgctggcgtcgcctcgcggtccgcagtggat
ggcgcctgaagcccgacacagtgagcttctcggttacgtggtccggtacggtaggcttgatcaacagcggcgagcttttgatcaacgaccttttggcttccct
ggagagcgcgagattctccgcgctgtagagtccagtagtcctcgtgcacgatccgctgttatccagctaagcgacaccatttgtgcacagcgcgattcgagatgccagcgcaat
gacattcttgcagtgtatctcagtgtccagcacgccagccacgatcgacgatctgctctcttgctgacaaaaagcaagagaacatagcgttgcctttggtaggtccagcggccgggaggaactc
tttgatccggtcctgaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgcccgactggccgatggagcgaaatgtagtgcttacgttgtcc
cgcattggtacgtggtgtacgcgccgcagtgtcgctgctgCgactgggcaatcgctgcgctgccgactcagctcacacttgagctggcag
gcttatcttggacaagatcgcttgcctgcgcgccgccggtagagattgttcactacgtgaagaatttgttgaagccgagctagagtagtcggcaatatgtctaacaatt
PstI

FIG. 3D

cgttcaagccgacgccgcttcgcggcgcggcttaactcaagcgttagaTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACG

PvuI

ATCAAGGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCcgatcgAGGATTTTTCGGCGCTGCGCTACGTCCGCkACCGCGTTGAGGGATC
AAGCCACAGCAGCCCACTCGACCTCTAGCCGACGAGCCAAGGGATCTTTTGGAATGCTGCTCCGTCAGCGTTTCCGACGTTTGGGTGGTTGAACAGAGAGTCATT
ATCGTACGGAATGCCAAGCACTCCCGAGGGGAACCCTGTGGTTGGCATGCCACATACAAATGGACGAACGGATAAACCTTTCACGCCCTTTAAATATCCGTTATTCTAATAAA

HindIII

CGCTCTTTTCTCTTAGGTTTACCCGCCCAATATATCCTGTCAAACACTGATAGTTTAAACTGAAGGGCGGGAAACGACAATCTgatccccATCA

FIG. 3E